Interaction of the Initiator Protein of an IncB Plasmid with Its Origin of DNA Replication

T. Betteridge, J. Yang, A. J. Pittard, and J. Praszkier*

Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Victoria 3010, Australia

Received 15 October 2002/Accepted 6 January 2003

The replication initiator protein RepA of the IncB plasmid pMU720 was purified and used in DNase I protection assays in vitro. RepA protected a 68-bp region of the origin of replication of pMU720. This region, which lies immediately downstream of the DnaA box, contains four copies of the sequence motif 5'AANCNG-CAA3'. Mutational analyses identified this sequence as the binding site specifically recognized by RepA (the RepA box). Binding of RepA to the RepA boxes was ordered and sequential, with the box closest to the DnaA binding site (box 1) occupied first and the most distant boxes (boxes 3 and 4) occupied last. However, only boxes 1, 2, and 4 were essential for origin activity, with box 3 playing a lesser role. Changing the spacing between box 1 and the other three boxes affected binding of RepA in vitro and origin activity in vivo, indicating that the RepA molecules bound to ori^B interact with one another.

Miniplasmid pMU720, a derivative of the large, low-copynumber, conjugative plasmid pMU707, is a member of incompatibility group B (9). Replication of pMU720 requires the synthesis of the initiator protein, RepA, which is rate limiting for replication. Studies of pMU720 and its close relative, the IncI₁ plasmid Collb-P9, have shown that translation of *repA* is coupled to translation of a leader peptide, RepB, and has to be activated by the formation of a pseudoknot immediately upstream of its Shine-Dalgarno sequence (2, 4, 5, 14, 27, 37). Expression of *repA*, and consequently the copy numbers of pMU720 and Collb-P9, is regulated by a small antisense RNA molecule, which on binding to its complementary sequence in the leader region of the *rep* mRNA not only preempts pseudoknot formation but also inhibits translation of the leader peptide (1–4, 6, 23, 31, 32, 37–39).

Initiation of replication of pMU720 and ColIb-P9 requires the presence in cis of two DNA sequences, the origin of replication (ori) and CIS (22, 25, 34). CIS, which lies between the repA coding sequence and ori, is composed of two domains. The repA-proximal domain has strong transcription termination activity and is thought to be involved in the loading of RepA onto ori (25). The repA-distal domain plays a role in positioning of sequences within ori at an appropriate distance and on the correct face of the DNA helix with respect to the repA-proximal domain of CIS (25). The ori of pMU720 contains the sequence 5'TTATCCACA3' (DnaA box), which is a consensus sequence for a binding site of the DnaA protein. Although this sequence is not essential for replication of pMU720, its deletion lowered the copy number of the wildtype plasmid threefold. Deletion of an additional 10 bp immediately downstream of the DnaA box inactivated ori (25).

The RepA protein acts in *cis*; i.e., it preferentially activates the *ori* of the DNA molecule that was the template for its mRNA. However, in the absence of cognate *ori* in *cis*, RepA is

able to activate its *ori* in *trans*, i.e., when it is present on a second plasmid (24). The level of *repA* expression required to initiate replication from an *ori* in *trans* was much higher than that required for an *ori* in *cis*, indicating that activation in *trans* is relatively inefficient (24). Moreover, when RepA acts in *trans*, *CIS* is not required for initiation of replication (24), which is consistent with the notion that the role of *CIS* is to tether the nascent RepA and load it onto the *ori* present in *cis* (25).

In this paper, we describe the purification of the RepA protein and characterize its binding sites in *ori*. We find that in vitro, purified RepA binds to the DnaA box-proximal region of *ori*. This binding is sequence specific and involves interactions with nucleotides in a sequence motif, 5'AANCNGCAA3' (RepA box), which is repeated four times. Binding of RepA to the RepA boxes is sequential, with the box closest to the DnaA binding site (box 1) occupied first and the most distant boxes (boxes 3 and 4) occupied last.

MATERIALS AND METHODS

Bacterial strains, plasmids, and phages. The strains of Escherichia coli K-12 used in this study are given below. JM101 [Δ (lac-proAB) supE thi F'(traD36 proA+B+ lacF2 Δ M15)] (20) was used for cloning and propagating M13 derivatives. XL1 Blue MRF' { Δ (mcrA)183 Δ (mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacF2 Δ M15 Tn10 (Tet^r)} (Stratagene) was used to grow M13 derivatives which had undergone mutagenesis as described by Vandeyar et al. (35). JP3438 (thr-1 leuB6 thi-1 lacY1 gal-351 supE44 tonA21 hsdR4 rpoB364 recA56) was used for propagating pMU720 derivatives and for all copy number determinations.

Bacteriophage vectors used to clone fragments for DNA sequencing and mutagenesis were M13tg130 and M13tg131 (15). The plasmids used are described in Table 1.

Media, enzymes, and chemicals. The minimal medium used was half-strength buffer 56 (21) supplemented with 0.2% glucose, thiamine (10 μg/ml), and necessary growth factors. Enzymes and chemicals of a suitable grade were purchased commercially and not purified further. [^{35}S]dATPαS (deoxyadenosine 5'-[α- ^{35}S]thiotriphosphate) (>1,000 Ci/mmol) for use in sequencing and [α- ^{32}P]dATP and [α- ^{32}P]dCTP (3,000Ci/mmol) for end labeling of DNA fragments were obtained from Amersham Biosciences Pty. Ltd. Ampicillin was used at a final concentration of 50 μg/ml, chloramphenicol was used at 10 μg/ml, kanamycin was used at 20 μg/ml, isopropylthiogalactoside (IPTG) was used at 1 mM, and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) was used at 25 μg/ml

^{*} Corresponding author. Mailing address: Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Victoria 3010, Australia. Phone: 61 3 8344 7751. Fax: 61 3 9347 1540. E-mail: judy@ariel.its.unimelb.edu.au.

TABLE 1. Plasmids

Plasmid	Relevant characteristics ^a	Reference or source
pMU720	Miniplasmid derived from pMU707; Gal IncB	9
pMU1585	pSU39 carrying nt 1–1916 of pMU720 with T642C and T640C mutations; Km; RepA plasmid	24
pMU1586	pAM34 carrying nt 1916–2170 of pMU720; Ap Cm lacI ^q ; ori ^B plasmid	24
pMU1599	pAM34 carrying nt 1–2170 of pMU720; Ap Cm lacI ^q ; IncB	25
pMU1600	pAM34 carrying nt 1916–2076 of pMU720; Ap Cm lacI ^q ; ori ^B plasmid	This study
pAM34	pMB1 derivative in which the preprimer RNA is expressed from $lacZpo$; $lacI^q$ Ap Sp	12
pSU39	p15A replicon; Km	8

^a Abbreviations: Ap, ampicillin resistance; Cm, chloramphenicol resistance; Km, kanamycin resistance; Sp, spectinomycin resistance. Mutations introduced into ori^B of pMU1600 are described in Results.

Recombinant DNA techniques. Plasmid and bacteriophage DNA were isolated and manipulated as described by Sambrook and Russell (28). DNA was sequenced by using a model 377 DNA sequencer and ABI Big Dye terminators (Perkin-Elmer Corporation) or by the method of Sanger et al. (29), which was modified in that T7 DNA polymerase was used instead of the Klenow fragment and terminated chains were uniformly labeled with [35 S]dATP α S. Oligonucleotide-directed in vitro mutagenesis reactions were performed on single-stranded M13 templates, using the method of Vandeyar et al. (35). Oligonucleotides were purchased from GeneWorks Ltd. DNA sequencing was used to screen for and confirm the presence of mutations.

Purification of the RepA protein. Plasmids that overproduce recombinant RepA protein carrying a hexahistidine-S or a thioredoxin-hexahistidine tag were constructed by using pET-30a(+) and pET-32a(+) vectors (Novagen Inc.), respectively. Recombinant RepA proteins were expressed in E. coli strain BL21(DE3) by IPTG induction. Cells were grown for 6 h at 37°C in a 2 ml of Luria broth supplemented with ampicillin at final concentration of 50 µg/ml. This culture was diluted 50-fold with fresh medium supplemented with 0.4% glucose, grown at 30°C to a cell density (A_{600}) of 1, and induced with 0.1 mM IPTG at room temperature for 2 h. Cells were harvested, washed with an equal volume of ice-cold 100 mM Tris-HCl (pH 8.3)-0.1 M NaCl-1 mM EDTA-0.1 mM phenylmethylsulfonyl fluoride, and resuspended in 1/10 volume of ice-cold phosphate buffer (phosphate-buffered saline [PBS] [27 mM NaCl, 0.54 mM KCl, 2 mM Na₂HPO₄, 0.4 mM KH₂PO₄, pH 8], 1% Triton X-100, 5 mM dithiothreitol [DTT], 1 mM phenylmethylsulfonyl fluoride, 1 mM EDTA, 5% glycerol). All subsequent steps were performed at 4°C. The cells were disrupted by sonication in a Braun Labsonic 2000 sonicator, and the pellet was collected by centrifugation and washed three times with phosphate buffer. The pellet, which was composed mainly of RepA fusion protein in the form of inclusion bodies, was resuspended in fresh phosphate buffer by slow rotation at 4°C for 2 h. Inclusion bodies were collected by centrifugation and resuspended in 0.6 ml of ice-cold glycine-NaOH (pH 10) by slow rotation for 2 h at 4°C. A soluble fraction of the RepA fusion protein, which was 80 to 90% pure and represented 20 to 50% of the total in inclusion bodies, was collected in the supernatant and stored at either -20 or -70°C in 50% glycerol. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to monitor the purification procedure. The protein concentration in the samples was estimated by protein assay or by SDS-PAGE.

Preparation of 3'-end-labeled DNA fragments. DNA fragments used for the DNase I footprinting experiments were prepared by digestion of the appropriate plasmid with a first enzyme (EcoRI for the upper strand and SaII for the lower strand) and end labeled (at the 3' end) by using Klenow fragment in the presence of $[\alpha^{-32}P]dATP$ or $[\alpha^{-32}P]dATP$ and $[\alpha^{-32}P]dCTP$ (3,000 Ci/mmol). Following phenol-chloroform extraction and ethanol precipitation, labeled DNA was digested with the second enzyme, and a fragment of 192 bp carrying the ori^{18} (ori of pMU720) sequence was purified in 5% polyacrylamide gel. The concentrations of labeled DNA preparations were determined by measuring their optical density at 260 nm.

DNase I footprinting experiments. The end-labeled $oni^{\rm B}$ fragment was mixed with 12.5 to 200 ng (240 to 3,800 fmol) of RepA fusion protein in a binding buffer (20 mM Tris-HCl [pH 8.3], 40 mM NaCl, 5 mM MgCl₂, 0.5 mM DTT, 0.5 mM CaCl₂, 1 mM ATP, and 0.6 mg of nuclease-free bovine serum albumin [BSA] per ml), equilibrated at room temperature for 10 min, and then incubated at 30°C for 20 min. A standard binding reaction mixture contained 10 ng of labeled DNA (160 fmol) in a total volume of 25 μ l, and RepA dilutions were prepared in ice-cold glycine-NaOH [pH 10] buffer. The protein-DNA complexes were digested with 0.0125 U of DNase I (Amersham Biosciences Pty. Ltd.) for 30 s at

room temperature, and 10 μ l of phenol and 15 μ l chloroform were added to stop the reaction. Samples were extracted and ethanol precipitated. The pellets were resuspended in formamide dye mix and analyzed on a 6% polyacrylamide sequencing gel. Following electrophoresis, the gel was scanned with a phosphorimager (Fuji FLA-3000G) and then exposed to Kodak-XAR film at -70°C for 48 h.

EMSA and DNase I footprinting in gel slice. The 32 P-labeled wild-type $ori^{\rm B}$ fragment (0.2 kb) used in the electrophoretic mobility shift assay (EMSA) and DNase I footprinting in gel slice experiments was generated as follows. The oligonucleotide primer TN97 (5'GTTCACAGTGGTTTCAGAGAT3') was labeled with 32 P at the 5' end by using $[\gamma^{-32}$ P]ATP and T4 polynucleotide kinase. The $ori^{\rm B}$ fragment was amplified by PCR with the 32 P-labeled primer NT97, primer TN96 (5'CCAGTGAATTGCTGCAGAGATC), and the M13tg130 derivative carrying the ori region of pMU720. The amplified ori fragment (labeled at the 5' end of the bottom strand) was then purified on a native polyacrylamide gel. The reaction mixtures for EMSA (25 μ l) contained approximately 3 nM end-labeled ori fragment, 8 to 480 nM RepA protein, 0.2 μ g of poly(dI-dC), and 15 μ g of BSA in DNase I footprinting buffer. The reaction mixtures were incubated at 30°C for 20 min before being analyzed on a 5% native polyacrylamide gel.

The DNase I in gel slice assay was performed by a procedure based on that described by Straney et al. (33). Gel slices containing various RepA-DNA complexes as well as the free DNA fragment from EMSA were excised from the polyacrylamide gel. The gel slices were each incubated in $10~\mu l$ of covering buffer (10 mM Tris-HCl [pH 8.0], 2 mM DTT, 5% glycerol, 0.5 mg of BSA per ml, and 0.4 U of DNase I per ml) (Amersham Biosciences Pty. Ltd.) at room temperature for 15 min. Five microliters of starting solution (50 mM MgCl $_2$ and 50 mM NaCl) was then added to each sample, and the reactions were continued for a further 2 min before being terminated by the addition of 30 μl of stop solution (100 mM EDTA and 2% SDS). DNA was extracted from the gel slices and analyzed on a sequencing gel.

Construction of plasmids for use in copy number determination. The two-plasmid system (24) was used to study in vivo the interactions of RepA with $ori^{\rm B}$ in trans. The RepA-producing plasmid was pMU1585 (24), and the ori plasmid was pMU1600 or its derivatives carrying mutations in the $ori^{\rm B}$ sequence. pMU1600 was constructed by replacing the 254-bp SacI-SacII $ori^{\rm B}$ fragment of pMU1586 (24) by the 161-bp SacII-SacI fragment carrying minimal $ori^{\rm B}$. Plasmid pMU1600 contains the modified pMB1 replicon from pAM34 (12), in which the essential preprimer RNA is transcribed from the lacZ promoter operator. Since this plasmid contains the $lacI^{\rm q}$ gene, replication of its pAM34 replicon requires the presence of a lac inducer, such as IPTG. Thus, in the absence of IPTG, replication of the ori plasmids is dependent on the RepA provided in trans by the RepA plasmid. The presence of a constitutively expressed chloramphenicol acetyltransferase (CAT) reporter gene allows estimations of the copy numbers of pMU1600 and its derivatives to be made.

Introduction of *ori* and RepA plasmids into *E. coli* cells. *ori* and RepA plasmids were cotransformed into *E. coli* K-12 strain JP3438 by the method of Chung et al. (11). Cells were plated onto medium containing half-strength buffer 56 (21), 0.2% glucose, 0.2% Casamino Acids, thiamine (10 μ g/ml), ampicillin, chloramphenicol, and kanamycin, with and without IPTG, and incubated for 72 h at 37° C. Plates were checked after 48 and 72 h of incubation, and the number and size of colonies produced in the presence and absence of IPTG were compared. Single colonies from plates without IPTG were used for copy number estimations.

Measurement of CAT activity. CAT activity of mid-log-phase cultures, grown in minimal medium containing 0.4% glucose, thiamine, leucine, threonine, kanamycin, ampicillin, and chloramphenicol, was assayed as described by Shaw (30).

Cells were disrupted by sonication with a Braun Labsonic 2000 sonicator, and cellular debris was removed by centrifugation prior to the assays. Each assay was performed at least six times. CAT activity was expressed as units per milligram of protein.

Protein assay. The concentration of protein in cleared cell lysates was determined by the method of Bradford (10), using BSA as a standard.

RESULTS

Purification of the RepA protein. RepA is a 343-amino-acid protein with a predicted molecular mass of 40.5 kDa (26). It is highly basic, due to the presence of 42 arginine residues and 24 lysine residues, which constitute 19.2% of the total amino acids, and has a predicted pK_i of 10.21. The RepA protein was produced in two fusion forms. One carried a small hexahistidine-S tag and the second carried a 109-amino-acid thioredoxin tag as well as a hexahistidine tag. Both recombinant proteins formed aggregates when overproduced in E. coli, with only a small fraction present in soluble form. Moreover, although recombinant RepA that had been solubilized under denaturing conditions with guanidine hydrochloride or urea could be purified by using His-bind resin, the protein formed aggregates again when the chaotropic agent was removed. It was therefore decided to purify RepA from its inclusion bodies (see Materials and Methods). The purified inclusion bodies were solubilized in glycine-NaOH buffer at pH 10. RepA protein present in the soluble fraction was active in EMSA and DNase I footprinting assays and remained active in the glycine-NaOH buffer (pH 10) at -20° C for at least 6 months. A typical yield was 0.5 to 2 mg of protein from a 100-ml culture.

The purification was monitored at each step by SDS-PAGE (Fig. 1). As seen in lane 7, the RepA protein recovered after treatment with glycine-NaOH was ~90% pure. The same procedure was successful for both RepA fusion proteins, and both were active in vitro, producing an identical DNase I footprint (data not shown). The thioredoxin-RepA fusion, hereafter referred to as RepA protein, was used in all subsequent experiments, as it gave better yields of the final product.

DNase I footprinting of ori^B fragment. In vivo studies of RepA acting in trans, when repA and ori of pMU720 were present on separate plasmids, showed that a sequence of 161 bp, corresponding to nucleotides (nt) 1916 to 2076 of pMU720, was sufficient to ensure efficient replication of the ori plasmid (J. Praszkier and J. Pittard, unpublished data). Therefore, this 161-bp fragment (Fig. 2), designated minimal ori^B, was used in DNase I footprinting experiments to identify binding sites for RepA. The footprints obtained (Fig. 3) show that a 68-bp region of ori^B, extending from nt 1928 to 1995, was protected from DNase I digestion upon binding of RepA. At the lowest protective concentration (30 nM), RepA protected nt 1928 to 1949 on the top strand of ori^B and enhanced cleavage at nt 1951 (Fig. 3, top strand, lane 3). Increasing the concentration of RepA resulted in extension of the protected region to nt 1972 and increased cutting at position 1951 (Fig. 3, top strand, lane 2). At the highest concentration of RepA tested (250 nM), the footprint extended to nt 1995, and there was enhanced cleavage at the hypersensitive site (Fig. 3, top strand, lane 1). Protection of the lower strand of ori^B appeared to be stronger than that of the top strand, as complete protection was seen with 30 nM RepA (Fig. 3, lower strand, lane 3); there was no extension of the protected region upon increase in the concentration of RepA. The protected region extended from nt 1928 to nt 1994, with hypersensitive sites corresponding to the adenosine residues at positions 1949 and 1950.

To confirm that RepA bound to sequences in ori^B in an ordered and successive manner, ori^B -RepA complexes were analyzed by EMSA. Four major complexes were observed with increasing amounts of RepA (Fig. 4A). The fastest migrating complex, complex I, was seen at the lowest ratio of RepA to ori^B fragment. As the amount of RepA increased, complexes of progressively slower mobility appeared, with a concomitant decrease in the amounts of the faster-migrating complexes. Analysis of complexes I and III by DNase I footprinting in gel slice (Fig. 4B) showed that complex I consisted of RepA bound only to region 1 (nt 1925 to 1953 on the bottom strand), whereas in complex III RepA occupied all three protected regions (nt 1925 to 1991 on the bottom strand). The hypersensitive sites at positions 1949 and 1950 were present in the footprint of complex III but not in that of complex I.

The shifting of the hypersensitive site, from position 1958 on the top strand of the free DNA to position 1951 in the RepAbound DNA, suggests there is a change in the DNA structure as a consequence of binding of RepA. The progressive extension of the RepA footprint on the top strand of ori^B and the progressive appearance of DNA-RepA complexes of decreasing electrophoretic mobility, observed with increasing concentrations of this protein, indicate the presence of at least three binding sites, with differing affinities for RepA, between nt 1928 and 1995 of ori^B .

Identification of RepA binding sites in *ori*^B. The region of *ori*^B that is protected from digestion by DNase I upon binding of RepA contains a sequence motif, 5'AANCNGCAA3', which is repeated four times in the top strand (Fig. 2). The first repeat (box 1) is located in the region protected at the lowest concentration of RepA (region 1), the second repeat (box 2) is in the region protected at intermediate concentrations of

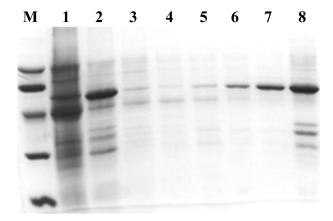


FIG. 1. Analysis by SDS-PAGE of fractions obtained during purification of the 52-kDa Trx-His-RepA fusion. Lane M, markers (from top to bottom, 97.4, 66, 45, 31, and 21.5 kDa); lane 1, supernatant fraction after induction with 0.1 mm IPTG; lane 2, cell pellet after induction with 0.1 mm IPTG; lane 3, supernatant from first PBS wash of the pellet; lane 4, supernatant from second PBS wash of the pellet; lane 5, supernatant from third PBS wash of the pellet, lane 6, soluble fraction after 10-min treatment of the pellet in glycine-NaOH buffer; lane 7, final soluble fraction; lane 8, inclusion bodies remaining in the pellet after treatment with glycine-NaOH buffer.

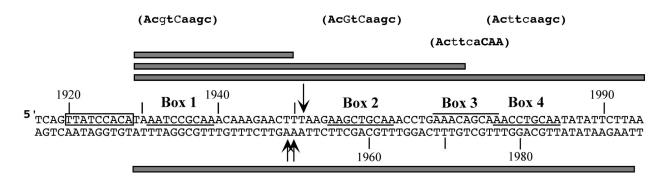




FIG. 2. Sequence of the $or^{\rm B}$ fragment. The DnaA box is boxed, and the repeated motif 5'AANCNGCAA3' (boxes 1 to 4) is indicated by lines below or above the sequence. The DNase I-hypersensitive sites are indicated by arrows. The regions protected by different concentrations of RepA from cleavage by DNase I are marked by the bars above the sequence. Sequences of scrambled RepA boxes 1 to 4 are shown in parentheses (nucleotides in boldface are conserved in all boxes, and the substitutions are shown in lowercase letters).

RepA (region 2), and repeats 3 and 4, which overlap by one base (boxes 3 and 4), are in the region that is protected only at high concentrations of RepA (region 3). To determine whether boxes 1 to 4 are involved in the binding of RepA to ori^B , the sequence of each was altered individually by oligonucleotide-directed mutagenesis, and the effects of these changes on binding of RepA in vitro and origin activity in vivo were examined. Each change involved multiple substitutions within the targeted box, with care being taken not to alter significantly the GC content of the sequence.

Changing the sequence of box 1 from 5'AATCCGCAA to 5'AcgtCaagc (nucleotides in uppercase and boldface are conserved in all boxes, and the substitutions are in lowercase) abolished protection of region 1 (nt 1928 to 1949) and resulted in loss of the hypersensitive site at nt 1951 (nt 1952 to 1995) (Fig. 5). Changing the sequence of box 2 from 5'AAGCTGCAA to 5'AcGtCaagc abolished protection of region 2 (nt 1952 to 1972) and resulted in loss of the hypersensitive site at nt 1951 but had no effect on protection of region 1 (Fig. 5). Changing the sequence of box 3 from 5'AAACAGCAA to 5'ActtcaCAA abolished protection of the part of region 3 (nt 1971 to 1975) that contained the base substitutions and increased accessibility to DNase I of position 1974, without affecting protection of regions 1, 2, and the rest of region 3 or the hypersensitive site at nt 1951 (Fig. 5). Changing the sequence of box 4 from 5' AACCTGCAA to AACtgcCAA abolished protection of the entire region 3 (nt 1972 to 1995) but did not affect protection of regions 1 and 2 or the hypersensitive site at nt 1951 (Fig. 5).

The effects of these (Table 2, mutations 1 to 4) and other mutations in boxes 1 to 4 on the replicative activity of $ori^{\rm B}$ in vivo were tested by replacing the ori fragment of pMU1600 with the mutant ori and determining the copy number of the resulting plasmid when RepA was provided in trans (24). As shown in Table 2, ori plasmids carrying mutated box 1 or box 2 were not able to replicate, whereas the plasmid carrying the

mutated box 4 had a copy number 100-fold lower than that of the wild type. Changing eight out of the nine bases of box 4, so that the sequence of the mutant box resembled those of box 1 and 2 mutants, resulted in complete loss of ori^B activity (Table 2, mutation 5). Furthermore, changing two out of seven conserved base pairs, an AT to CG change at position 2 and CG to GC change at position 4 in RepA box 1 or 4, was sufficient to reduce the copy number of the ori^B plasmid more than 10-fold, and change of the same bases in box 2 reduced the copy number 3-fold (Table 2). By contrast, mutation of box 3 had no significant effect on the copy number of the *ori* plasmid. However, replacement of eight out of the nine bases of box 3, which also changed the first base of box 4 from A to C, reduced the copy number of the ori^B plasmid 5-fold (Table 2, mutation 6), whereas this A-to-C substitution alone reduced it 1.7-fold (Table 2, mutation 7), suggesting that substitutions in box 3 were also contributing to the defect in replication. These data confirm the critical importance of boxes 1, 2, and 4 to the interaction of RepA with ori^B and indicate that box 3 may play a lesser role.

The sequence of the region separating boxes 1 and 2 is not important for replication. The effect of base substitution in the 15-bp region separating RepA boxes 1 and 2 on the activity of ori^B in vivo was examined. This region was targeted because it contains the sites of enhanced sensitivity to DNase I digestion that are the consequence of RepA binding. Two or three bases at a time were changed. Since this region is AT rich (11 out of 15 residues are AT), some of the changes involved replacement of A/T by C/G to determine whether disruption of the poly(AT) tract would affect the efficiency of replication.

None of the substitutions in the intervening region between boxes 1 and 2, including replacement of the A/T residues that are the site of enhanced sensitivity to DNase I by C/G residues, had a significant effect on efficiency of replication (Table 3). Moreover, replacement of T1951 did not cause the loss of the

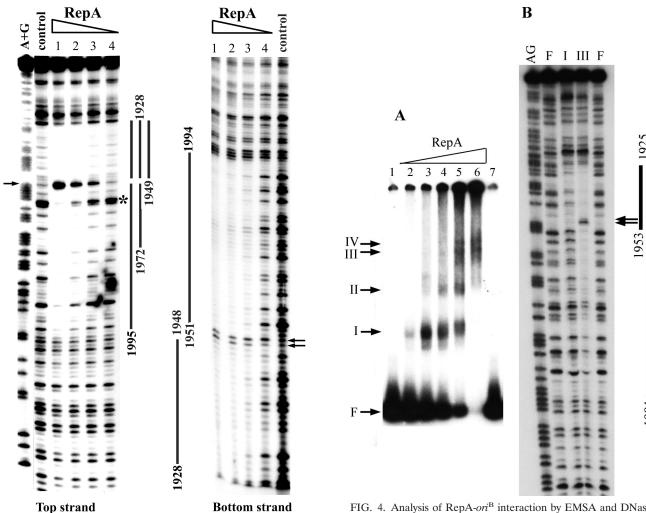


FIG. 3. DNase I footprinting of wild-type $ori^{\rm B}$. A 20 nM concentration of the 3'-end-labeled (either in the top or in the bottom strand) double-stranded $ori^{\rm B}$ fragment was incubated with 15 to 250 nM RepA protein and subjected to partial digestion with DNase I. Regions of DNA protected from digestion by DNase I are indicated by vertical lines, with the position of the last band protected indicated. Lanes: control, no RepA; 1, 250 nM RepA; 2, 125 nM RepA; 3, 30 nM RepA; 4, 15 nM RepA; G+A, Maxam-Gilbert sequencing reaction. Arrows indicate sites of enhanced cleavage by DNase I; the asterisk marks position 1958.

hypersensitive band at that position (data not shown). Replacement of three bases downstream of box 2 had no effect on replication, whereas replacement of three bases within the box reduced the efficiency of replication. These data show that the sequence of the intervening region between the first two boxes of $ori^{\rm B}$ is not important for replication.

The spacing between the RepA boxes of *ori*^B is critical for replication. To determine whether positioning of the RepA boxes with respect to one another is important for binding of RepA in vitro and replication in vivo, bases were deleted or inserted in the region separating boxes 1 and 2.

Deletion of two bases (AA at position 1946 [Fig. 2]) resulted in loss of the hypersensitive site at position 1951 (Fig. 6) in vitro and in a fivefold reduction of copy number in vivo (Table

FIG. 4. Analysis of RepA-ori^B interaction by EMSA and DNase I footprinting in gel slice. (A) EMSA was carried out as described in Materials and Methods, using ori^B fragment labeled at the 5' end of the bottom strand. Samples in lanes 1 to 7 contained 0, 8, 16, 32, 160, 480, and 0 nM RepA protein, respectively. The positions of free DNA (F) and various complexes (I to IV) are marked. (B) Free DNA and complexes I and III from EMSA (see panel A) were analyzed by DNase I footprinting in gel slice as described in Materials and Methods. The regions protected by RepA from digestion by DNase I are indicated by vertical lines, with the position of the last band protected indicated, and the hypersensitive bands are marked with arrows. Lanes: AG, AG ladder generated by Maxam-Gilbert sequencing; F, free DNA digested with DNase I; I, complex I digested with DNase I; III, complex III digested with DNase I.

4). Insertion between boxes 1 and 2, at position 1952, of five bases, which displaces the boxes by half a turn of the helix, or of nine bases, which moves them apart by almost a full turn of the helix, resulted in loss of the hypersensitive site at position 1951 (data not shown) in vitro and in a complete loss of *ori*^B activity in vivo (Table 4).

These data show that effective binding of RepA to $ori^{\rm B}$ in vitro and origin activation in vivo require that RepA box 2 and/or box 4 be at the wild-type distance from box 1. Neither shortening this distance nor increasing it is permissible, even when helical phasing is only slightly altered.

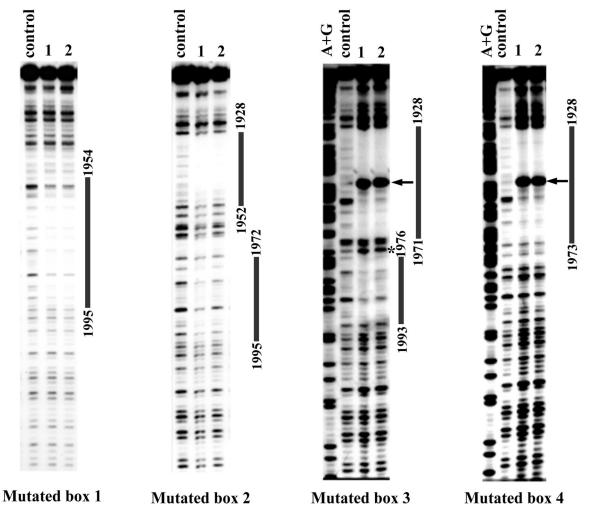


FIG. 5. DNase I footprinting of ori^B fragment carrying mutations in boxes 1 to 4. A 20 nM concentration of the 3'-end-labeled top strand of each ort^B fragment was incubated with 125 or 250 nM RepA protein and subjected to partial digestion with DNase I. Regions of DNA protected from digestion by DNase I are indicated by vertical lines, with the position of the last band protected indicated. Lanes: control, no RepA; 1, 250 nM RepA; 2, 125 nM RepA; G+A, Maxam-Gilbert sequencing reactions. Arrows indicate sites of enhanced cleavage by DNase I.

TABLE 2. Effects of mutations in RepA boxes 1 to 4 on the copy number of the ori^B plasmid

Mutation no.	Location of mutation(s)	Change in sequence ^a	Copy number of <i>ori</i> plasmid relative to wild type ^b
1	Box 1	AATCCGCAA to AcgtCaagc	NR^c
2	Box 2	AAGCTGCAA to AcGtcaagc	NR
3	Box 3	AAACAGCAAACCTGCAA to ActtcaCAAACCTGCAA	1.0
4	Box 4	AAACAGCAAACCTGCAA to AAACAGCAAACtgcCAA	0.01
5	Box 4	AAACAGCAAACCTGCAA to AAACAGCAActtcaagc	NR
6	Boxes 3 and 4	AAACAGCAAACCTGCAA to ActtcaagcACCTGCAA	0.2
7	Box 4	AAACAGCAAACCTGCAA to AAACAGCAcACCTGCAA	0.6
8	Box 1	AATCCGCAA to AcagCGCAA	0.06
9	Box 2	AAGCTGCAA to AcagTGCAA	0.3
10	Box 3	AAACAGCAA to AcAgAGCAA	1.0
11	Box 4	AACCTGCAA to AcagTGCAA	0.09

^a Base substitutions are shown in lowercase letters, and residues conserved in the RepA box motif are in boldface; the underlined residues are shared by boxes 3 and 4. b Copy number was determined by assaying the CAT activity of *E. coli* strain JP3438 carrying the *ori*^B plasmids, in the absence of IPTG, and expressed relative to that of the plasmid carrying wild-type *ori*^B. The values shown are averages from at least four independent determinations.

c NR, *ori*^B plasmid was unable to replicate in the absence of IPTG.

TABLE 3. Effects of mutations in the spacer region separating RepA boxes 1 and 2, and in the 3' flank of box 2, on the
copy number of the $ori^{\rm B}$ plasmid

Location of mutation(s)	Change in sequence ^a	Copy number of <i>ori</i> plasmid relative to wild type ^b
Wild type	AATCCGCAAACAAAGAACTTTAAGAAGCTGCAAACCTG	1.0
Between boxes 1 and 2	AATCCGCAAACAgcGAACTTTAAGAAGCTGCAAACCTG	1.0
Between boxes 1 and 2	AATCCGCAAACAAActtCTTTAAGAAGCTGCAAACCTG	1.0
Between boxes 1 and 2	AATCCGCAAACAAAGAACccTAAGAAGCTGCAAACCTG	1.1
Between boxes 1 and 2	AATCCGCAA ACAAAGAACTcatAGAAGCTGCAAACCTG	1.0
Between boxes 1 and 2	AATCCGCAAACAAAGAACTTTccGAAGCTGCAAACCTG	0.9
Box 2	AATCCGCAAACAAAGAACTTTAAGAcagTGCAAACCTG	0.3
Box 2	AATCCGCAAACAAAGAACTTTAAGAAGCTGCcAgCCTG	0.6
Box 2	AATCCGCAAACAAAGAACTTTAAGAAGCTGCtttCCTG	0.6
3' flank of box 2	AATCCGCAAACAAAGAACTTTAAGAAGCTGCAAAACgac	0.9

^a Base substitutions are shown in lowercase letters, and residues conserved in the RepA box motif, which is underlined, are in boldface.

DISCUSSION

The first step in initiation of replication of pMU720 is thought to involve binding of the replication initiator protein RepA to specific sequences in ori. This notion is based on the finding that in vivo RepA showed specificity for its cognate ori (25). However, there has been no physical evidence to show that RepA, or any of the Rep proteins related to it, binds to the region designated ori. We have overexpressed RepA of pMU720 and developed a simple purification procedure, producing RepA that was active in vitro. This purified RepA was shown to protect a 68-bp region of ori^B (nt 1928 to 1995) from digestion by DNase I. RepA showed a hierarchical binding pattern, preferentially protecting a 22-bp region (nt 1928 to 1949) adjoining the DnaA box, which marks the 5' end of ori^B. With progressive increases in the concentration of RepA, the protected region extended to nt 1971 and then to nt 1995. Mutational analyses identified the sequence 5'AANCNGCA A3' (the RepA box) as the binding site specifically recognized by RepA. This sequence occurs four times within the region protected by RepA, but only boxes 1, 2, and 4 are essential for origin activity.

The RepA protein of pMU720 shows 39.4% sequence identity with RepA of the IncL/M plasmid pMU604 (7) and 32.3% identity with RepA of plasmid pSW800 (40). Examination of the designated ori regions of pMU604 and pSW800 revealed that both contain three copies of a RepA box-like motif, whose positions correspond to those of boxes 1, 2, and 4 except that the distance between boxes 1 and 2 of pSW800 was 1 bp shorter than in the other two plasmids (Fig. 7). The consensus of the pMU604 motif is 5'NANCYGCAA3' and that of pSW800 is 5'TACCCGCAA3', with bases shown in boldface present in the motifs of all three plasmids. The conservation of both the sequence and the location of the RepA boxes in the ori of pMU604 and pSW800 strongly suggests that they too are the binding sites for their cognate RepA proteins. The only sequences other than the RepA boxes that are conserved in all three oris are the DnaA boxes, located at the 5' ends of these oris, and a 6-mer, 5'TCTTAA3', which is present 5 nt downstream of RepA box 4 (Fig. 7). The importance of this 6-mer for *ori* activity is being investigated.

The finding that the positioning of the RepA boxes is critical

for origin activity is indicative of interaction between the RepA protein molecules bound to ori^B . This notion is supported by the data showing that the appearance of the hypersensitive site at position 1951 on the top strand, which is indicative of a

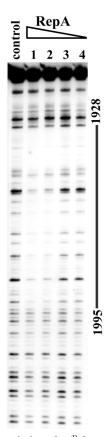


FIG. 6. DNase I footprinting of *ori*^B fragment deleted of 2 bp (nt 1948 and 1949) in the region separating RepA boxes 1 and 2. A 20 nM concentration of the 3'-end-labeled top strand of the *ori*^B fragment was incubated with 15 to 250 nM RepA protein and subjected to partial digestion with DNase I. The region of DNA protected from digestion by DNase I is indicated by a vertical line, with the position of the last band protected indicated. Lanes: control, no RepA; 1, 250 nM RepA; 2, 125 nM RepA; 3, 30 nM RepA; 4, 15 nM RepA;

^b Copy number was determined by assaying the CAT activity of *E. coli* strain JP3438 carrying the *ori*^B plasmids, in the absence of IPTG, and expressed relative to that of the plasmid carrying-wild type *ori*^B. The values shown are averages from at least four independent determinations.

TABLE 4. Effects of deletion or insertions in the spacer region separating RepA boxes 1 and 2 on the copy number of the ori^B plasmid

Mutation	Change in sequence ^a	Copy number of <i>ori</i> plasmid relative to wild type ^b
Wild type	AATCCGCAAACAAAGAACTTTAAGAAGCTGCAAACCTG	1.0
Deletion 2	AATCCGCAAACAAAG—CTTTAAGAAGCTGCAAACCTG	0.2
Insertion 5	AATCCGCAAACAAAGAACTTTaagacAAGAAGCTGCAAACCTG	NR
Insertion 9	AATCCGCAAACAAAGAACTTTaagacgtctAAGAAGCTGCAAACCTG	NR

^a Bases inserted are shown in lowercase letters, and residues conserved in the RepA box motif, which is underlined, are in boldface.

^c NR, $or^{\dot{B}}$ plasmid was unable to replicate in the absence of IPTG.

change in the local conformation of DNA, required that both box 1 and box 2 be occupied by RepA. However, RepA box 1 is not spaced at full helical turns from boxes 2 and 4, as is usual in cases of interactions between protein molecules, where it is important that their binding sites be on the same face of the DNA helix. Moreover, deletion of 2 bp, which improves the spacing between the corresponding residues of boxes 1 and 2 to two helical turns, resulted in a fivefold reduction in replicative activity in vivo (Table 4) and an aberrant binding in vitro. This defect is unlikely to be due to loss of the two adenosine residues deleted in this mutant, as their replacement by thymidine residues had no effect on replication (Table 3). Insertion between boxes 1 and 2 of 5 bp, which moves them half a helical turn apart, or of 9 bp, which moves them almost a full helical turn apart and improves the helical spacing between them, inactivated the origin. The simplest explanation for these seemingly anomalous findings is that binding of RepA to boxes 1 and 2 alters the local conformation of the spacer DNA separating them, bringing boxes 2 and 4 into helical phase with box 1.

IncB plasmids are distantly related to IncFII plasmids and share with them features such as the general organization of the replicon, the use of antisense RNA to control copy number via the regulation of expression of *rep*, lack of an absolute requirement for DnaA, and the preferential *cis* activity of the Rep protein. However, there is no significant sequence homology between the replicons of these two groups of plasmids. The Rep protein of the IncFII plasmid R1 has been purified and

used in vitro (13, 18). DNase I footprints were similar to those described here for pMU720 in that the primary binding site was located at the 5' end of ori, immediately downstream of the DnaA box, and in that at high concentrations of Rep, protection extended in the 3' direction to produce a footprint of \sim 90 bp (13, 18). There was evidence of cooperative binding of multiple Rep molecules, but the binding sites were not identified (13). It was proposed that Rep binds in a sequencespecific manner to only two sites, one at each end of the protected region, and that protein-protein interaction between these bound Rep molecules leads to looping of the intervening DNA, facilitating sequence-independent filling of the loop by additional molecules of Rep (13). This is unlike the binding of RepA to ori^B, where there are four specific binding sites for RepA within the 68-bp protected region and binding proceeds sequentially, with the site closest to the DnaA box occupied first and the and two distal sites occupied last.

Binding of DnaA, the replication initiator protein of *E. coli*, to the five DnaA boxes (R1 to R4 and M) in its origin of replication (*oriC*) also proceeds in an ordered, sequential, and sequence-specific manner, with R4 occupied first and M occupied last (17, 36). Moreover, the positioning and orientation of the leftmost box, R1, is critical, as insertion of 10 bp between it and box M inactivated the origin (19). This is despite the fact that these two boxes lie 46 bp apart, a much greater distance than that separating RepA boxes 1 and 2. However, *oriC* appears to be less sensitive to point mutations than *ori*^B, as changes in a single DnaA box, whether it was R1, R2, R3, or



FIG. 7. Alignment, using Clustal W, of *ori*^B (nt 1920 to 2096) of pMU720, with the corresponding sequences of pMU604 and pSW800. Nucleotides conserved in all three sequences are indicated by asterisks below the sequence. The DnaA boxes are underlined, and RepA boxes 1, 2, and 4 are indicated by bars above the sequence.

^b Copy number was determined by assaying the CAT activity of *E. coli* strain JP3438 carrying the *ori*^B plasmids, in the absence of IPTG, and expressed relative to that of the plasmid carrying wild-type *ori*^B. The values shown are averages from at least four independent determinations.

R4, had little effect on origin activity (16). On the other hand, scrambling the sequence of R1, R2, R4, or M, so that it bore no resemblance to that of the consensus DnaA box, inactivated *oriC*-dependent replication in a wild-type host (16). Scrambling the sequence of R3 had no effect on *oriC* activity, but point mutations in R3 augmented the effect of point mutations in the other DnaA boxes (16), a situation resembling that of mutants of RepA box 3. It has been proposed that R3 plays a regulatory role in the initiation of replication of the *E. coli* chromosome (16).

In $E.\ coli$, the binding of DnaA to the 9-mer DnaA boxes in oriC results in destabilization of the AT-rich region located to the left of R1, leading to the unwinding of the DNA double helix in that region. Although DnaA is not absolutely required for the replication of pMU720 (L. Borrell, J. Yang, A. J. Pittard, and J. Praszkier, unpublished data), it contributes to efficient replication of this plasmid, as deletion of the DnaA box in ori^B resulted in a \sim 3-fold reduction in the plasmid copy number (24, 25). However, this defect in replication could be partly compensated for by increasing the expression of repA (25). These data suggest that RepA, like DnaA, is involved in promoting open complex formation. The observation that binding of RepA to ori^B leads to a change in the local conformation of DNA supports this notion.

ACKNOWLEDGMENT

This work was supported by a grant from the National Health and Medical Research Council.

REFERENCES

- Asano, K., C. Hama, S.-I. Inoue, H. Moriwaki, and K. Mizobuchi. 1999. The plasmid Collb-P9 antisense Inc RNA controls expression of the RepZ replication protein and its positive regulator *repY* with different mechanisms. J. Biol. Chem. 274:17924–17933.
- Asano, K., A. Kato, H. Moriwaki, C. Hama, K. Shiba, and K. Mizobuchi. 1991. Positive and negative regulations of plasmid Collb-P9 repZ gene expression at the translational level. J. Biol. Chem. 266:3774–3781.
- Asano, K., and K. Mizobuchi. 1998. Copy number control of IncIa plasmid ColIb-P9 by competition between pseudoknot formation and antisense RNA binding at a specific RNA site. EMBO J. 17:5201–5213.
- Asano, K., and K. Mizobuchi. 1998. An RNA pseudoknot as the molecular switch for translation of the repZ gene encoding the replication initiator of IncIa plasmid ColIb-P9. J. Biol. Chem. 273:11815–11825.
- Asano, K., H. Moriwaki, and K. Mizobuchi. 1991. An induced mRNA secondary structure enhances repZ translation in plasmid Collb-P9. J. Biol. Chem. 266:24549–24556.
- Asano, K., T. Niimi, S. Yokoyama, and K. Mizobuchi. 1998. Structural basis for binding of the plasmid Collb-P9 antisense Inc RNA to its target RNA with the 5'-rUUGGCG-3' motif in the loop sequence. J. Biol. Chem. 19: 11826–11838.
- Athanasopoulos, V., J. Praszkier, and A. J. Pittard. 1995. The replication of an IncL/M plasmid is subject to antisense control. J. Bacteriol. 177:4730–4741.
- Bartolomé, B., Y. Jubete, E. Martínez, and F. de la Cruz. 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. Gene 102:75–78.
- Bird, P. I., and J. Pittard. 1983. Demonstration of a third incompatibility function on plasmids already incompatible with group P and group I plasmids. Plasmid 9:191–200.
- Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248–254.
- Chung, C. T., S. Niemela, and R. H. Miller. 1989. One step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc. Natl. Acad. Sci. USA 86:2172–2175.
- Gil, D., and J.-P. Bouché. 1991. ColE1-type vectors with fully repressible replication. Gene 105:17–22.

- Giraldo, R., and R. Díaz. 1992. Differential binding of wild type and a mutant RepA protein to *oriR* sequence suggests a model for initiation of plasmid R1 replication. J. Mol. Biol. 228:787–802.
- Hama, C., T. Takizawa, H. Moriwaki, and K. Mizobuchi. 1990. Role of leader peptide synthesis in repZ gene expression of the Collb-P9 plasmid. J. Biol. Chem. 265:10666–10673.
- Kieny, M. P., R. Lathe, and J. P. Lecocq. 1983. New versatile cloning and sequencing vectors based on bacteriophage M13. Gene 26:91–99.
- Langer, U., S. Richter, A. Roth, C. Weigel, and W. Messer. 1996. A comprehensive set of DnaA-box mutations in the replication origin, oriC, of Escherichia coli. Mol. Microbiol. 21:301–311.
- Margulies, C., and J. M. Kaguni. 1996. Ordered and sequential binding of DnaA protein to oriC, the chromosomal origin of Escherichia coli. J. Biol. Chem. 271:17035–17040.
- Masai, H., and K. Arai. 1987. RepA and DnaA proteins are required for initiation of R1 plasmid replication in vitro and interact with the oriR sequence. Proc. Natl. Acad. Sci. USA 84:4781–4785.
- Messer, W., H. Hartmann-Kuhlein, U. Langer, E. Mahlow, A. Roth, S. Schaper, B. Urmoneit, and B. Woelker. 1992. The complex for replication initiation of Escherichia coli. Chromosoma 102:S1-S6.
- 20. Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. 101:20-78.
- Monod, J., G. Cohen-Bazire, and M. Cohen. 1951. Sur la biosynthèse de la b-galactosidase (lactase) chez *Escherichia coli*. La specificité de l'induction. Biochim. Biophys. Acta 7:585–599.
- Mori, A., K. Ito, K. Mizobuchi, and Y. Nakamura. 1995. A transcription terminator signal necessary for plasmid Collb-P9 replication. Mol. Microbiol. 17:291–301.
- Praszkier, J., P. Bird, S. Nikoletti, and A. J. Pittard. 1989. Role of countertranscript RNA in the copy number control system of an IncB miniplasmid. J. Bacteriol. 171:5056–5064.
- Praszkier, J., S. Murthy, and A. J. Pittard. 2000. Effect of CIS on activity in trans of the replication initiator protein of an IncB plasmid. J. Bacteriol. 182:3972–3980.
- Praszkier, J., and A. J. Pittard. 1999. Role of CIS in replication of an IncB plasmid. J. Bacteriol. 181:2765–2772.
- Praszkier, J., T. Wei, K. Siemering, and J. Pittard. 1991. Comparative analysis of the replication regions of the IncB, IncK, and IncZ plasmids. J. Bacteriol. 173:2393–2397.
- Praszkier, J., I. W. Wilson, and A. J. Pittard. 1992. Mutations affecting translational coupling between the *rep* genes of an IncB miniplasmid. J. Bacteriol. 174:2376–2383.
- Sambrook, J., and D. W. Russell. 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
- Shaw, V. W. 1975. Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. Methods Enzymol. 43:737–755.
- Siemering, K. R., J. Praszkier, and A. J. Pittard. 1993. Interaction between the antisense and target RNAs involved in the regulation of IncB plasmid replication. J. Bacteriol. 175:2895–2906.
- Siemering, K. R., J. Praszkier, and A. J. Pittard. 1994. Mechanism of binding
 of the antisense and target RNAs involved in the regulation of IncB plasmid
 replication. J. Bacteriol. 176:2677–2688.
- Straney, D. C., S. B. Straney, and D. M. Crothers. 1989. Synergy between Escherichia coli CAP protein and RNA polymerase in the lac promoter open complex. J. Mol. Biol. 206:41–57.
- Tanaka, K., H. Sakai, Y. Honda, T. Nakamura, A. Higashi, and T. Komano. 1992. Structural and functional features of *cis*-acting sequences in the basic replicon of plasmid Collb-P9. Nucleic Acids Res. 20:2705–2710.
- Vandeyar, M. A., M. P. Weiner, C. J. Hutton, and C. A. Batt. 1988. A simple and rapid method for the selection of oligodeoxynucleotide-directed mutants. Gene 65:129–133.
- Weigel, C., A. Schmidt, B. Ruckert, R. Lurz, and W. Messer. 1997. DnaA protein binding to individual DnaA boxes in the Escherichia coli replication origin, oriC. EMBO J. 16:6574–6583.
- Wilson, I. W., J. Praszkier, and A. J. Pittard. 1993. Mutations affecting pseudoknot control of the replication of B group plasmids. J. Bacteriol. 175:6476–6483.
- Wilson, I. W., J. Praszkier, and A. J. Pittard. 1994. Molecular analysis of RNAI control of *repB* translation in IncB plasmids. J. Bacteriol. 176:6497–6508.
- Wilson, I. W., K. R. Siemering, J. Praszkier, and A. J. Pittard. 1997. Importance of structural differences between complementary RNA molecules to control of replication of an IncB plasmid. J. Bacteriol. 179:742–753.
- Wu, C. Y., J. F. Fu, and S. T. Liu. 2001. The replicon of pSW800 from Pantoea stewartii. Microbiology 147:2757–2767.